CLAIMS:

1. A method for including a document in an index in a hyperlinked environment, comprising the acts of:

receiving a document to be processed;

locating a set of documents that include hyperlinks to the document;

retrieving anchortext associated with at least one of the hyperlinks;

parsing the anchortext into one or more tokens;

for each token:

determining a weight for the token,

determining whether the weight assigned to the token exceeds a threshold token weight; and

indexing the document under the token, if the token weight assigned to the token exceeds the threshold token weight.

2. The method of claim 1, wherein the indexing act comprises including in the index an indication of weight for each token under which each page is indexed.

3. The method of claim 1, wherein the indexing act comprises assigning to the token a location within the index corresponding to part of the page being indexed that is allocated for tokens having a higher degree of importance than other tokens in the same page.

4. The method of claim 3, wherein the indexing act comprises assigning to the token a location within the index that corresponds to the beginning of the page being indexed.

5. The method of claim 1, wherein the weight of each token is based on its frequency of occurrence within the index.

6. The method of 1, wherein the act of determining a weight comprises:

determining a first frequency at which the anchortext appears in the index;

determining a second frequency at which each token derived from the anchortext appears in the index; and

assigning a weight to the token, wherein the weight is a function of the first and second frequencies.

7. The method of claim 6, further comprising dividing the first frequency by the second frequency to produce a weight quotient; and

multiplying the weight quotient by an anchor text count for the token.

8. The method of 6 further comprising determining a normalized weight for each token.

9. A program product embedded in a machine-readable medium for including a document in an index in a hyperlinked environment, comprising the instructions for:

receiving a document to be processed;

locating a set of documents that include hyperlinks to the document;

retrieving anchortext associated with each hyperlink;

parsing the anchortext into one or more tokens;

and program instructions for each token comprising instructions for:

determining a weight for the token,

determining whether the weight assigned to the token exceeds a threshold token weight; and

indexing the document under the token, if the token weight assigned to the token exceeds the threshold token weight.

10. The computer program product of claim 9 wherein the indexing instruction comprises including in the index an indication of weight for each token under which each page is indexed.

11. The computer program product of claim 9, wherein the weight of each token is based on its frequency of occurrence within the index.

12. The computer program product of claim 9, wherein the indexing act comprises assigning to the token a location within the index that corresponds to the beginning of the page being indexed.

13. The computer program product of claim 9, wherein the weight of each token is based on its frequency of occurrence within the index.

14. The computer program product of claim 9, wherein the instruction of determining a weight comprises:

determining a first frequency at which the anchortext appears in the index;

determining a second frequency at which each token derived from the anchortext appears in the index; and

assigning a weight to the token, wherein the weight is a function of the first and second frequencies.

15. The program product of 13, further comprising the instruction of determining a normalized weight for each token.

16. A system for indexing a document in a hyperlinked environment, comprising:

a receiver for receiving a document to be processed for inclusion in an index of documents;

a module for locating a set of documents that include hyperlinks to the document;

a module for retrieving anchortext associated with each hyperlink;

a parsing module for parsing the anchortext into one or more tokens; and

a module for:

determining a weight for the token,

determining whether the weight assigned to the token exceeds a threshold token weight; and

indexing the document under the token, if the token weight assigned to the token exceeds the threshold token weight.